Universal properties of culture

Alexandru-Ionuţ Băbeanu, Leandros Talman, and Diego Garlaschelli

Lorentz Institute for Theoretical Physics, Leiden University, The Netherlands

(Dated: February 20, 2017)

Understanding the formation of subjective human traits, such as preference and opinions, is an important, but poorly explored problem. An essential aspect is that traits collectively evolve under the repeated action of social influence interactions, which is the focus of many quantitative studies of cultural dynamics. In this paradigm, dynamical models require that all traits are fixed when specifying the "initial cultural state". Typically, this initial state is randomly generated, from a uniform distribution over the set of possible combinations of traits. However, recent work has shown that the outcome of social influence dynamics strongly depends on the nature of the initial state: if this is sampled from empirical data instead of being generated in a uniformly random way, a higher level of cultural diversity is found after long-term dynamics, for the same level of propensity towards collective behavior in the short-term; moreover, if the initial state is obtained by shuffling the empirical traits among people, the level of long-term cultural diversity is in-between those obtained for the empirical and random counterparts. The current study repeats the analysis for multiple empirical data sets, showing that the results are remarkably similar, although the matrix of correlations between cultural variables clearly differs across data sets. This points towards robust structural properties inherent in empirical cultural states, likely due to universal laws governing the dynamics of culture in the real world. The analysis suggests, first, that this dynamics operates close to criticality and second, that it is driven by more than just social influence, implications which were not recognized previously.

I. INTRODUCTION

A fundamental challenge for the social sciences consists of understanding the formation (or origin) of preferences, opinions, values and beliefs – shortly, of "cultural traits". This is intimately linked with understanding human decision making, thus having strong relevance for disciplines like sociology, economics or political science. As will become more apparent below, trait formation also has implications on the faith of cultural diversity and on collective human behavior, problems that have recently attracted much attention from natural scientists. However, the problem of cultural trait formation appears to be poorly explored. Almost thirty years ago, Ref. [1] was arguing that questions of how preferences form tend to either be avoided or receive answers that are only valid for certain cultures or societies. Currently, some efforts are being made to address such questions [2–4], driven by the economic relevance of preferences. Arguably, a complete answer to the the problem of trait formation can only be given by a general, formal theory of trait dynamics, which would account for both the dynamical nature of culture (the fact that traits evolve with time) and the multidimensional nature of culture (the fact that, at the level of every individual, cultural traits cannot exist independently of other cultural traits).

Formulating and validating such a theory is beyond the purpose of this study, although this work does provide certain guidelines in this direction. Nonetheless, it is instructive to reflect on the types of processes that conceivably drive the dynamics of cultural traits in the real world, which can be seen as potential ingredients of a general theory. First, a change of an individual trait may happen by copying the trait of another individual as a result of an interaction – a "social influence" mechanism. Second, a change of an individual trait may happen when real-world experience contradicts an upheld personal view of a certain issue – a "reality-based consistency" mechanism. Third, a change may be a consequence of the person reconsidering his/her position with respect to a certain issue, due to some inconsistency or tension with his/her position on some other issue – a "cross-topic consistency" mechanism. Most likely, there are other relevant ingredients that one may think of, likely having to do with how the human mind works, but these are enough for illustrating the complexity of the problem of trait formation – obviously, these mechanisms would manifest in parallel and would be coupled to each other at the level of every individual. For the rest of this manuscript, it is convenient to use the term "exogenous dynamics" to refer to trait dynamics driven by social influence. By contrast, "endogenous dynamics" is used to jointly refer to all other, potentially relevant ingredients which may be defined independently of social influence, such as the reality-based and cross-topic consistency mechanisms mentioned above – this "endogenous–exogenous" distinction should not be confused with that made in Ref. [2], which uses the same terminology for a different purpose, although within a similar context.

The exogenous dynamics of traits has recently received much attention from a highly interdisciplinary research community, with significant input from physics [5], which lead to the development of various models and quantitative tools. In addition to embracing the dynamical nature of traits, this approach often also embraces the multidimensional nature of culture, in which case it referred to as "cultural dynamics" – in contrast to similar research focusing on single-dimensional dynamics, in which case it is referred to as "opinion dynamics". This work mainly

builds on previous work in "cultural dynamics", for which the so-called Axelrod model [6] is very representative. In this paradigm, an individual is understood as a long sequence of cultural traits, or a "cultural vector", which evolves in time. Every entry of the vector corresponds to one dimension of culture, also referred to as cultural variable or "cultural feature". There are many versions of the Axelrod model present in the literature, each incorporating a different mix of ingredients along with social influence [7–15]. As far as the authors are aware, neither of these versions includes endogenous processes in its dynamical rules – but see the discussion concerning "cultural drift" in Sec. V A. The need of bringing together exogenous and endogenous ingredients has also been recently signaled by Ref. [16] (using a different terminology), which is also proposing a first model which achieves such an integration. Unlike Ref. [16], the present work does not attempt an exogenous-endogenous integration on the theory side. Instead, it studies the joint manifestation of endogenous and exogenous processes in available empirical data, which consists of partial snapshots of the real world dynamics, as explained below.

One might argue that the effects of endogenous processes are actually accounted for when specifying the initial conditions of exogenous dynamics models, namely the set of initial cultural vectors – the "initial cultural state". However, the initial conditions are usually specified randomly, using a uniform probability distribution over the set of possible cultural vectors – a uniform distribution in "cultural space". As an exception, Ref. [17] is using a non-uniform distribution, but one that can be factorized as a product between Poisson, feature-level distributions, which does not entail any correlations. An apparent, a-posteriori justification of the uniform approach is that the main outcomes of exogenous dynamics seem to be largely independent of the specific realization of the initial cultural state, but dependent on the dynamical rules. However, this is tautological, as it is only valid for the class of cultural states that are representative of the uniform probability distribution. Moreover, it is difficult to believe that uniform randomness is a good description of the outcome of realistic endogenous processes – either alone or coupled to exogenous processes taking place "before the initial condition".

In fact, Refs. [18] and [19] showed that the initial conditions actually have a strong contribution in determining the outcome of exogenous dynamics: an initial cultural state constructed from empirical data seems to induce significant deviations with respect to the uniformly random case. On one hand, this suggests that there is much to be understood about the role of endogenous processes in preference formation and its interplay with exogenous processes. On the other hand, this implies that exogenous dynamics models can be used for probing the structure of the initial cultural state. Thus, if the cultural vectors in the initial state correspond to real individuals, the outcome of exogenous dynamics can be used as a quantitative tool for gaining insight about how real

individuals are distributed in cultural space, and indirectly about trait formation in the real world. This is, to a great extent, the direction of the research presented here.

A set of cultural vectors constructed from empirical data, also referred to below as an "empirical cultural state", is how the notion of "culture" is formalized in this study. It should be understood as a partial snapshot of cultural dynamics taking place in the real world: the set of cultural vectors corresponds to a random subset o individuals from a society, whose cultural traits are recorded within a relatively short period of time – typically via a social survey. Obviously, an empirical cultural state encodes all effects of trait formation – thus of both endogenous and exogenous processes – taking place in that social system before the snapshot. This is implicit for all the analysis and modeling of empirical cultural states done throughout this study. Separating the effects of endogenous processes from those of exogenous processes is something inherently difficult [1], which is not attempted at this point.

It is desirable that knowledge gained about trait formation in the real world has a "universal" nature, namely that the knowledge is valid across societies and across time. Consequently, studying the empirical cultural state is done with the aim of unraveling universal structural properties, that are independent of the particular sociocultural system from which the data is obtained. These properties could then be interpreted as signatures of general laws governing trait formation and used to indirectly achieve knowledge about these laws. The need of finding such generic properties, if they exists, is one of the essential motivations for a physics-inspired research approach. One can indeed wonder whether, just like uniformly random initial conditions lead to certain universal outcomes, different classes of initial conditions, in particular those constructed from empirical data, lead to different, but still universal outcomes. In other words, one would like to identify universality classes of cultural states, and in particular check whether empirical states associated to different data sources fall within the same universality class. The current work does not yet provide a rigorous way of identifying these universality classes in a statistical physics sense. However, it does emphasize remarkable statistical similarities between different data sets. These similarities suggest that different empirical cultural states might indeed belong to the same universality class. Formally, this is done using a certain mathematical tool which actually relies on a model of exogenous dynamics.

This tool, inspired by Ref. [18], is effectively a 2-dimensional plot which shows, for any given initial cul-

---

[1] Even the actual person might often not be able to tell what exactly caused a specific change of one of his/her traits, or whether it occurred "exogenously" or "endogenously" – in most cases both aspects would likely contribute in an intricate way, while parts of the combined process would take place subconsciously.

tural state, how a measure of long-term cultural diversity (LTCD) depends on a measure of short-term collective behavior (STCB) – or short-term social coordination. The LTCD quantity estimates the extent to which cultural diversity survives after a long period of exogenous trait dynamics governed by consensus-favoring social influence, in the absence of any other process. This is done using an Axelrod-type model [6] of trait dynamics. The STCB quantity estimates the propensity of the initial state to collective behavior on the short-term. This is done using a modification of the Cont-Bouchaud model [20] of social coordination. While LTCD is a property of the final state, STCB is a property of the initial state of the exogenous dynamics model. However, the definition of the STCB does employ, in a more implicit way, the idea of single-opinion dynamics driven by social influence, but taking place on a much shorter time-scale. In fact, as described in Sec. III, both the LTCD and the STCB quantities are functions of the same free parameter, the bounded confidence threshold ω, which controls the maximal distance in cultural space for which social influence can operate. The common dependence on this parameter is what allows for LTCD to be expressed as a function of STCB.

Ref. [18] showed that STCB and LTCD are mutually exclusive when the initial cultural state is generated in an uniformly random way. In other words, for uniformly random data, either the STCB or the LTCD attains a close-to-minimal value for any value of the bounded confidence threshold, apparently calling for a more complicated description or otherwise suggesting a paradox, since real-world societies seem to allow for both short-term collective behavior and long-term cultural diversity. However, when the initial cultural state is constructed from empirical data, the two aspects become clearly more compatible, with both quantities attaining intermediate values for a certain region of the threshold. This appears a parsimonious way of bringing about the expected coexistence of LTCD and STCB. Besides the uniformly random and empirical cultural states, a third class of initial conditions can be constructed, by shuffling the empirical traits among vectors, thus retaining only part of the information in the empirical data. This third class was found to entail a level of compatibility of LTCD and STCB which is intermediate between those obtained with empirical and uniformly random data. The current study is dedicated to checking the robustness of the LTCD-STCB behavior across different empirical data sets. As shown below, this behavior is indeed universal.

The rest of this manuscript is organized as follows. Sec. II rigorously explains the notion of "cultural state", by starting from the related concepts of "cultural space", "cultural feature", "cultural trait", "cultural distance" and "cultural vector". It also explains what it means to "shuffle" and to "randomize" a set of cultural vectors (SCV). Then, Sec. III describes two important quantities, associated to the notions of "long-term cultural diversity" (LTCD) and "short-term collective behavior" (STCB), as well as how these are combined to provide a quantitative tool that can emphasize the structure implicit in an empirical cultural state [18]. Finally, Sec. IV shows that this structure is robust across data sets and across geographical regions. Sec. V gives detailed discussions of results presented throughout the study, of possible criticism and of questions that can be further investigated. The manuscript is concluded in Sec. VI. Note that, although most notions explained in Sec. II and Sec. III have already been used in previous work, they are worth explaining at length here, due to their systematic usage with multiple data sets, as well as for clarifying various aspects that were previously only implicit.

II. THE FORMAL REPRESENTATION OF CULTURE

To understand how a cultural state is formally encoded here, it is worth starting with the notion of "cultural space". This originates from the development and study of models of exogenous trait dynamics, in particular of Axelrod-type models [6]. In this paradigm, one deals with a set of variables, called "cultural features", which encode information about various properties that individuals can have, properties that are inherently subjective and that can change under the action of "social influence" arising during person-to-person interactions. By construction, these variables are allowed to attain only specific values which are here called "cultural traits". The interpretation here is that cultural traits encode "preferences", "opinions", "values" and "beliefs" that people can have on various topics, where each topic is associated to one feature.

A cultural space consists of the set of all possible combinations of cultural traits entailed by the set of chosen cultural features, together with a measure of dissimilarity between any two combinations. Moreover, this dissimilarity, also called the "cultural distance", is defined in such a way that it satisfies all the properties of a metric distance (non-negativity, identity of indiscernibles, symmetry and triangle inequality). The so-called "Hamming" distance is commonly employed for this purpose, which is meaningful as long as there is no obvious ordering of the traits of any feature. A cultural space is thus an abstract, discrete, metric space, where each point corresponds to a specific combination of traits. However, the cultural space is mathematically not a vector space, since there is no notion of additivity attached to it.

A cultural state is essentially the selection of points in the cultural space that needs to be specified for the initial state of models of exogenous dynamics. Such a selection is also referred to here as a "set of cultural vectors" (SCV), where one "cultural vector" is one possible combination of traits. Formally, this is not a set in the rigorous sense, but a multiset, since it may contain duplicate elements – identical sequences of traits. However, duplicate

elements will rarely occur in the initial states constructed for this study, since the number of cultural vectors is in practice much smaller than the number of possible points the cultural space. Throughout this study, the notion of SCV is used interchangeably with that of "cultural state".

At this point, it may also be convenient to reflect on the more abstract notion of "cultural space distribution" (CSD), as a discrete probability mass function taking the cultural space as its support. If the SCV is constructed in a uniformly random way, one implicitly assumes that the cultural space distribution is constant, namely that all combinations of traits are equally likely. If, however, the SCV is constructed from empirical data, the inherent structure may be thought to correspond to non-homogeneities in an underlying CSD, for which the data is representative. Explicitly measuring or working with such an object is impractical for all purposes, at least for the present study, which is mainly due to the large number of points, the high a-priori dimensionality characterizing a typical cultural space and the sparseness of a sampling associated to an empirical set of cultural vectors. Nonetheless, the CSD, as a concept, may be useful, for instance when developing stochastic models for generating realistic SCVs – some steps in this direction have been taken in Ref. [19].

Of great importance for this study are the empirical SCVs, which are mainly constructed from social survey data. Cultural features are obtained from the questions that are asked in the survey, where the traits of each feature correspond to the answers that can be given to the respective question. Thus, a cultural vector consists of a sequence of answers that one individual has given to the list of questions in the survey. Importantly, a question is selected and encoded as a feature only if it is reasonably subjective, meaning that it does not ask about demographic or physical aspects concerning the individual, like place of residence, marital status, age, weight, and that every allowed answer should be plausible at least from a certain perspective of looking at the question, or for people with a certain background or a certain way of thinking. Moreover, a question is disregarded if the survey is defined in such a way that its list of a-priori allowed answers depends on what answers are given to other questions. All features remaining after this filtering – more details are provided in Sec. A of the Appendix – are assumed to contribute equally to the cultural distance, but the way they contribute depends on whether they are treated as nominal or as an ordinal variables. Specifically, the cultural distance d_{ij} between two vectors i and j is computed according to:

$$d_{ij} = 1 - \frac{1}{F} \sum_{k=1}^{F} \left[f_{\text{nom}}^k \delta(x_i^k, x_j^k) + (1 - f_{\text{nom}}^k) \frac{|x_i^k - x_j^k|}{q^k - 1} \right] = \frac{1}{F} \sum_{k=1}^{F} d_{ij}^k, \tag{1}$$

where F is the number of cultural features with k iterating over them, f_{nom}^k is a binary variable encoding the type of feature k (1 for nominal and 0 for ordinal), q^k is the range (number of traits) of feature k, $\delta(a, b)$ is a Kroneker delta function of traits a and b (of the same feature) and x_i^k is the trait of cultural vector i with respect to feature k. This definition reduces to the Hamming distance in case there are only nominal variables present. The second equality sign gives a formulation of the cultural distance as a sum over feature-level cultural distance contributions d_{ij}^k/F.

These feature-level contributions allow one to formulate, following Ref. [18], a notion of feature-feature covariance:

$$\sigma^{k,l} = \frac{\langle d_{ij}^k d_{ij}^l \rangle_{i,j \in \overline{1,N}}^{i<j} - \langle d_{ij}^k \rangle_{i,j \in \overline{1,N}}^{i<j} \langle d_{ij}^l \rangle_{i,j \in \overline{1,N}}^{i<j}}{F^2} \tag{2}$$

valid for any two features k and l, regarldess of f_{nom}^k and f_{nom}^l. Note that the averaging is performed over all $N(N-1)/2$ distinct pairs $(i, j), i \neq j$ of cultural vectors, rather than over all N cultural vectors. The feature-feature covariances can be used to define the feature-

feature (Pearson) correlation:

$$\rho^{k,l} = \frac{\sigma^{k,l}}{\sqrt{\sigma^{k,k} \sigma^{l,l}}} \tag{3}$$

which measures the extent to which large/small distances in terms of feature k are associated to large/small distances in terms of feature l. One can definitely see the $F \times F$ correlation matrix ρ as a reflection of a CSD that is compatible with the data. In general, however, the correlation matrix will only retain part of the information encoded in the CSD, first because $\rho^{k,l}$ retains only part of the information in the 2-dimensional contingency table of features k and l, second because a CSD is essentially an F-dimensional contingency table, which might imply all kinds of higher-order correlations.

Assuming the definition of cultural distance given by Eq. (1), a cultural space is already specified by the list of features taken from an empirical data set, together with the associated ranges and types. In this empirically-defined cultural space, it is meaningful to talk about several types of SCVs. First, an empirical SCV is constructed from the empirical sequences of traits of the individuals selected from those sampled by the survey. Second, a shuffled SCV is constructed by randomly permut-

ing the empirical traits among individuals, independently for every feature. Third, a random SCV is constructed by randomly choosing the trait of every person, for every feature. Note that the shuffled SCV exactly reproduces, for each feature, the empirical frequency of each trait, while disregarding all information about the frequencies of co-occurrence of various combinations of traits of two or more different features. Thus, shuffling destroys all feature-feature correlations $\rho^{k,l}$, as well as any higher-order correlations entailed by the empirical SCV, retaining only the information encoded in the marginal probability distributions associated to individual features. On the other hand, a random SCV retains nothing of the information inherent in the empirical SCV.

Finally, note that the mathematical definition of cultural distance illustrated by Eq. (1), already used in Refs. [18] in [19], is not unique nor very sophisticated. Other definitions might capture differences in opinions, preferences, values, beliefs, attitudes and associated behavior tendencies in better, more precise ways – see Ref. [21] for a sophisticated approach. However, the current definition is arguably good enough for the problems explored in this study and for how they are attacked. Since this research approach – using exogenous trait dynamics for probing the structure of cultural states – is only in its infancy, keeping all quantitative tools and models as simple as possible is desirable. Moreover, the current definition of cultural distance constitutes a good bridge between exogenous dynamics models and available social survey data.

III. LONG-TERM CULTURAL DIVERSITY AND SHORT-TERM COLLECTIVE BEHAVIOR

This section focuses on two quantities that are evaluated on sets of cultural vectors, namely the LTCD and STCB quantities mentioned above, which are defined in the same way as in Ref. [18]. Each of these quantities is based, explicitly or implicitly, on the idea of trait dynamics driven by social influence in a population of interacting agents. Each agent is associated, for the initial conditions, to one of the cultural vectors in the SCV that is studied. For simplicity, both quantities assume that there is no physical space nor a social network that would constrain the interactions between agents. In both cases, the interactions are assumed to only be constrained by how the agents are distributed in the cultural space. Specifically, only if the distance between two cultural vectors is smaller than the bounded confidence threshold ω are the two agents able to influence each other's opinions in favor of local consensus. This effectively says that there needs to be enough similarity between the cultural traits of two people if any of them is to convince the other of anything. This picture is inspired by assimilation-contrast theory [22], Ref. [10] being the first study that explicitly uses the bounded-confidence threshold in the context of trait dynamics, after having already been in use in the context

of opinion dynamics for some time – see Ref. [23] for an overview. The bounded confidence threshold ω functions like a free parameter on which both the LTCD and the STCB quantities depend, for any given SCV.

The LTCD quantity is a measure of the extent to which the given SCV favors cultural diversity on the long term, namely a survival of differences in cultural traits at the macro level, in spite of repeated, consensus-favoring interactions at the micro level. In the real world, boundaries between populations belonging to different cultures appear to be resilient with respect to social interactions across them [24–26]. The measure relies on a Axelrod-type model [6] of cultural evolution with bounded confidence, which is applied on the SCV. This is meant to computationally simulate the evolution of cultural traits under the action of dyadic social influence, in the absence of other processes that would be present in reality. According to this model, at each moment in time, two agents are randomly chosen for an interaction. If the distance between their cultural vectors is smaller than the threshold ω, then, with a probability proportional to $1 - \omega$, for one of the features that distinguishes between the two vectors, one of the agents changes its trait to match the other. With time, agents become more similar to those that are within a distance ω in the cultural space. The dynamics stops when several groups are formed, within which agents are completely identical to each other, but too dissimilar to those in other groups for any trait-changing interaction to occur. These groups are called "cultural domains", term formulated in the context of the original Axelrod model [6], which also included a physical/geographical, 2-dimensional lattice but no (explicit) bounded confidence threshold. The normalized number of such cultural domains for a given value of ω, averaged over multiple runs of the model, defines the LTCD quantity:

$$\text{LTCD}(\omega) = \frac{\langle N_D \rangle_\omega}{N}, \qquad (4)$$

where N_D is the cultural domains in the final (or absorbing) state of this model, the normalization being made with respect to N, the size of the SCV.

The STCB quantity is a measure of the extent to which the given SCV favors collective behavior (or social coordination) on the short term, namely the extent to which the agents associated to the cultural vectors in the set would tend to take actions or make choices in a similar, coordinated way rather than independently from each other. Bursts of fashion and popularity[27–29], rapid diffusion of rumors, gossips and habits[5, 30] and speculative bubbles and herding behavior on the stock markets[20, 31] are real-world examples of collective behavior on the short term. The measure relies on a Cont-Bouchaud type model [20], which deals with an aggregate choice or opinion of the entire agent population on one issue, which for simplicity is assumed here to be represented by a binary variable, such as liking vs disliking an item. According to the model, when collectively confronted with this issue,

the agents within a connected group effectively make the same choice or express the same opinion. In this context (where physical space and social network are disregarded), a connected group is a subset of agents that form a connected component in the graph obtained by introducing a link for every pair (i, j) of agents that are culturally close enough to socially influence each other $d_{ij} < \omega$. Based on this approximation, the aggregate, normalized choice of the entire population is expressed as a weighted average over the choices of the connected components, where the weight of the A'th component is the size S_A of this component. However, the group choices themselves are still assumed to be binary, equiprobable random variables with values $\{-1, +1\}$. Thus, the aggregate, normalized choice is also a random variable, but one that is non-uniformly distributed over some set of rational numbers within $[-1, 1]$, in a manner that depends on the set of group sizes $\{S_A\}_\omega$ induced by a specific value of the ω threshold. The spread of this aggregate probability distribution provides the coordination measure that defines the STCB. It turns out that this quantity can be analytically computed, for a given ω, according to [18]:

$$\text{STCB}(\omega) = \sqrt{\sum_A \left(\frac{S_A}{N}\right)^2_\omega}, \qquad (5)$$

where the summation is carried over the cultural connected components labeled by different A values. Note that only the sizes S_A of the components enter the calculation, which are in turn determined by the cultural graph obtained by applying the ω threshold to the set of pairwise cultural distance contributions d_{ij} between the cultural vectors. As visible in the formula, the STCB is higher when the agents are more concentrated in fewer and larger cultural connected components.

There is a crucial difference between the ideas behind the LTCD and the STCB measures: while the former assumes that agents move in cultural space under the action of social influence, the latter assumes that the agents remain fixed in cultural space while they make their decision on one issue with which they are just being confronted. Although, for computing the STCB, the reasoning in terms of connected cultural components implicitly assumes that social influence occurs within these components, this influence is supposedly too superficial and too short-lived too also alter the cultural vectors themselves. Thus, as their names suggest, the LTCD and STCB quantities are concerned with two different timescales: a long time-scale for which cultural vectors and distances are dynamic and a short time-scale for which cultural vectors and distances are fixed. Moreover, while LTCD requires computer simulations, the STCB is computed in an analytical way, although it does require an algorithmic step of finding the connected components in the cultural graph that is obtained by thresholding the matrix of cultural distances. Thus, LTCD can be seen as a characteristic of the final cultural state resulting from a long, exogenous dynamics process, while the STCB can

be can be seen as a property of the initial cultural state. Nonetheless, they both take as input the set of cultural vectors, and, as shown in Ref. [18], the combination of the two quantities turns out to be sensitive to the structure encoded in the set of cultural vectors.

It is worth explicitly illustrating, with Fig. 1, the behavior of the LTCD and the STCB quantities for a random SCV. The SCV is defined with respect to the cultural space of one of the data sets introduced in Sec. IV. Figs. 1(a) and 1(b) show, respectively, the dependence of the LTCD and STCB measures on the bounded-confidence threshold ω, while Fig. 1(c) shows the correspondence between the LTCD and STCB measures obtained by eliminating ω. The same data points are used for all 3 plots, where each point records all the 3 quantities (LTCD, STCB and ω). The LTCD quantity is averaged, for each point, over 10 runs of the exogenous dynamics model, with the associated standard deviations shown by the error bars.

Fig. 1(a) shows that LTCD is decreasing with ω: in the limit of large N, LTCD goes from 1 to 0 as ω goes from 0 to 1. This can be understood by keeping in mind that ω controls the range of interaction between agents in the cultural space. In general, convergence of agents happens in parallel in several regions of the cultural space, towards several points that are out of range of each other. The interaction range ω controls the expected number of such convergence points, which in turn determines the expected number of cultural domains in the final state and thus the LTCD value – the latter three quantities decrease with increasing ω. If ω is small enough, there is effectively no successful interaction and thus no movement in cultural space, so each agent "converges" to one, distinct point (assuming that all vectors are different from each other in the initial state). If ω is large enough, all agents tend to converge to the same point in the cultural space. Note that, in terms of ω, these two extreme cases are actually two regimes, separated by a sharp decrease of LTCD over some intermediate ω interval. This sharp decrease can actually be understood as an order-disorder phase transition, where the disordered phase corresponds to low ω, while the ordered phase corresponds to high ω. This transition has been previously studied in the context of the Axelrod model [14, 17], although in terms of a differently defined control parameter – the (average) feature range q rather than the bounded-confidence threshold ω.

Fig. 1(b) shows that STCB is decreasing with ω: in the limit of large N, STCB goes from 0 to 1 as ω goes from 0 to 1. This can be understood by keeping in mind that ω controls the extent to which agents are culturally connected to each other. Higher ω implies fewer, but larger connected components in the cultural graph, thus a higher predisposition for coordination. If ω is small enough, there is one connected component for every agent, while if ω is small enough, there is one connected component containing all agents. Similarly to above, these two cases correspond to two regimes separated by a sharp increase of STCB, which can be again understood
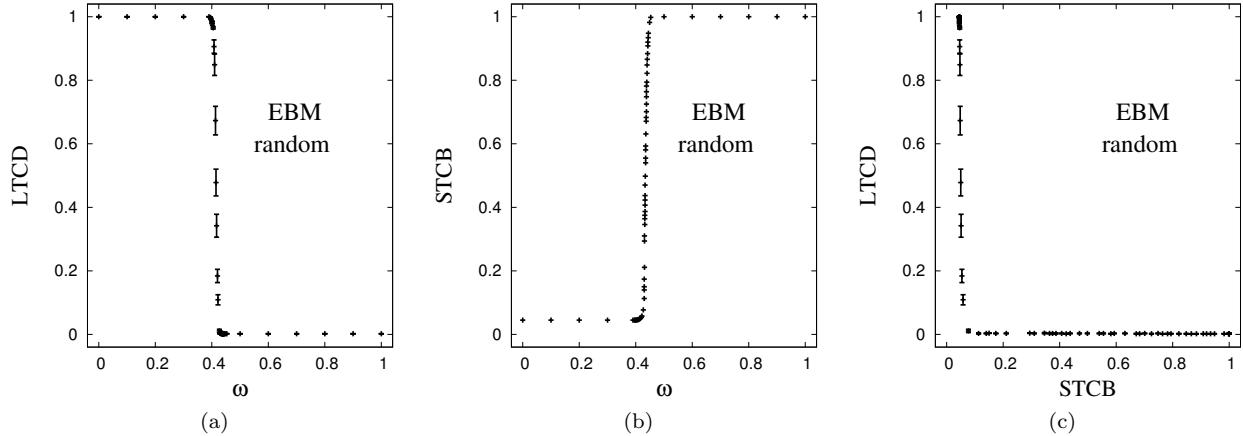
FIG. 1. Dependence of long-term cultural diversity and of short-term collective behavior on the bounded confidence threshold and their incompatibility, for a random set of cultural vectors. Showing the LTDC(ω) dependence (a), the STCB(ω) dependence (b) and the ω-induced LTDC-STCB correspondence (c), for a random set of cultural vectors of $N = 500$, in the cultural space of the Eurobarometer (EBM) data set (see Sec. IV).

as a phase transition – it is actually a symmetry breaking phase transition, as explained in Ref. [18].

Fig. 1(c) shows that, as ω increases, one goes from the upper-left corner (high LTCD, low STCB) to the lower-right corner (low LTCD, high STCB), by first passing through the lower-left corner (low LTCD, low STCB). In other words, the sharp decrease of LTCD happens before the sharp increase of STCB, meaning that the critical ω of the LTCD phase transition is lower than that of the STCB phase transition. This is also visible at a close, comparative inspection of Figs. 1(a) and 1(b). The ω-region for which both the LTCD and the STCB attain low values corresponds to a special situation for which there is a relatively high level of convergence in the final cultural state (low LTCD), in spite of a relatively low level of connectivity in the initial cultural state (low STCB). This is apparently explained by the fact that movement in cultural space at a certain moment during the exogenous dynamics simulation facilitates further movement that would not have been possible at an earlier moment, so it is enough to have a few pairs of agents that can initially influence each other to gradually set a large fraction of the other agents in motion and in the end achieve a large amount of convergence. In any case, Fig. 1(c) shows that at least one of the two quantities has to attain a close-to-minimal value, regardless of the bounded-confidence threshold ω.

According to the considerations above, long-term cultural diversity and short-term collective behavior seem to be mutually exclusive, suggesting a paradox [18], at least if one accepts that real socio-cultural systems allow for both aspects. However, the above calculations make use of a random SCV, meaning that the entire argument relies on the assumption that the underlying cultural space distribution is uniform. In fact, Ref. [18] showed that an empirical SCV allows for much more compatibility, with

both quantities attaining intermediate values for a certain ω interval – as shown in Sec. IV, this translates to a higher LTCD-STCB curve than the one shown in Fig. 1(c) – meaning that the apparent paradox is solved by using realistic data about cultural traits. Moreover, a shuffled SCV entails a compatibility level that is in between those entailed by a random and by an empirical SCV. Thus, Ref [18] showed that an empirical SCV has enough structure to dramatically affect the behavior of exogenous dynamics acting upon it, aspect which had been neglected in the past.

IV. UNIVERSALITY OF STRUCTURAL PROPERTIES

The findings of Ref. [18] are based on one data set. It is important to understand whether the observed properties are in fact robust across different populations and across different topics. This is accomplished by repeating the analysis of Ref. [18] on four data sets. These are taken from different sources, thus containing different cultural features and recording the traits of different people. The four data sources are: the Eurobarometer (EBM), containing opinions on science, technology and various European policy issues of people in EU countries [32]; the General Social Survey (GSS), containing opinions on a great variety of topics of people in the US [33]; the Religious Landscape (RL), containing religious beliefs and attitudes on certain political issues of people in the US [34]; Jester, containing online ratings of jokes [35].

Fig. 2 shows that the properties highlighted by the LTCD-STCB curves are indeed universal. The 4 panels correspond to the 4 empirical data sets that are used. In each panel, the 3 curves correspond to the 3 levels of preserving the empirical information: full information
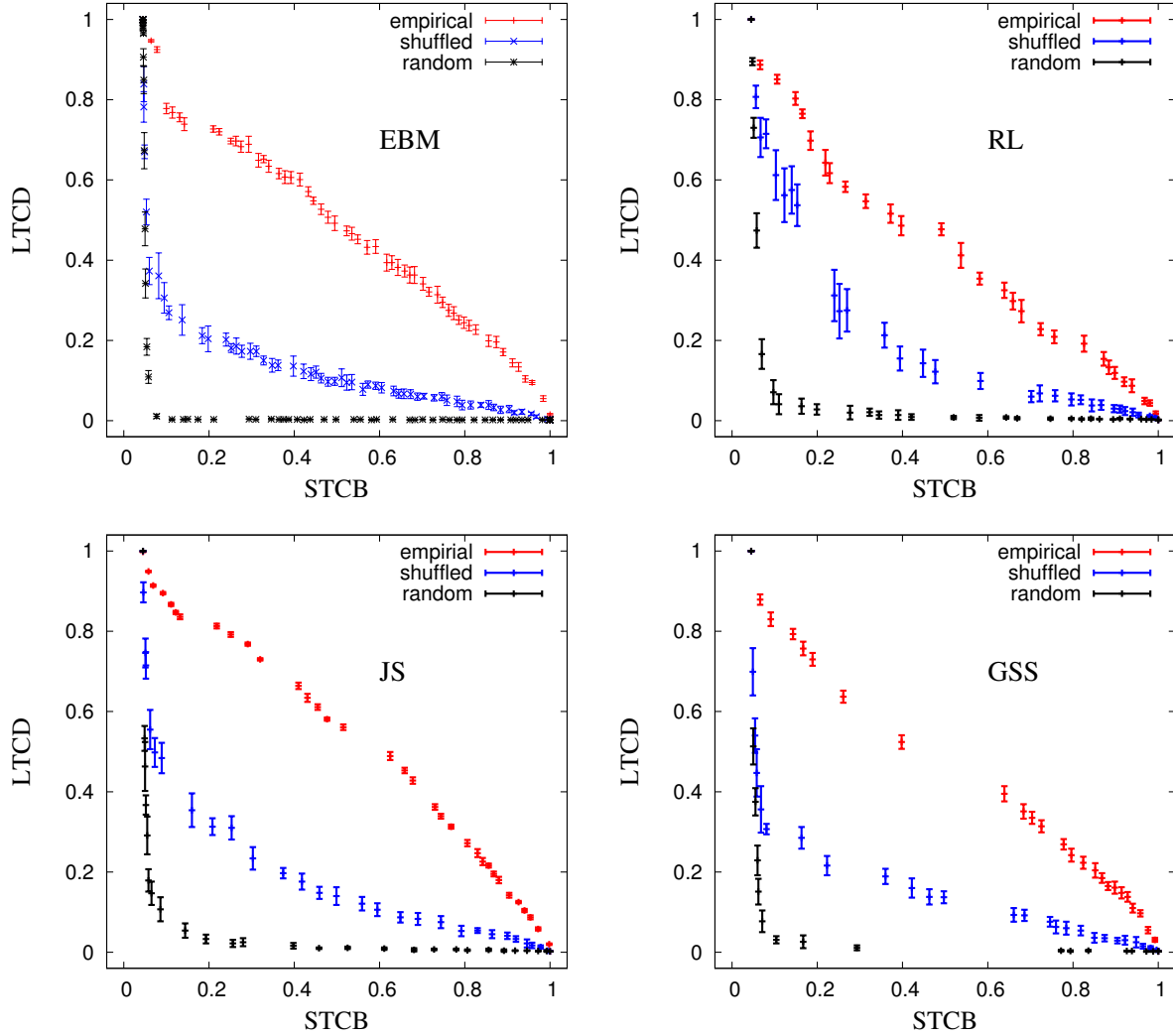
FIG. 2. The correspondence between long-term cultural diversity (LTCD) and short-term collective behavior (STCB) for the empirical (red), shuffled (blue) and random (black) sets of cultural vectors, for four data sets: Eurobarometer (EBM), General Social Survey (GSS), Religious Landscape (RL) and Jester (JS). Error bars denote standard deviations over multiple exogenous dynamics runs. There are $N = 500$ elements in each set of cultural vectors.

(red), corresponding to the empirical SCV; partial information (blue), corresponding to the shuffled SCV; no information (black), corresponding to the random SCV. Note that, for every data set, the empirical SCV allows for more compatibility between LTCD and STCB than the shuffled SCV, which in turn allows for more compatibility than the random SCV. Also note that the empirical LTCD-STCB correspondence is always close to the second diagonal. These qualitative observations constitute the basis for the claim of there being universal structural properties underlying empirical sets of cultural vectors.

In relation to aspects discussed at the end of Sec. II, the raising of the LTCD-STCB curve from the random to the shuffled and further to the empirical CSV visible in Fig. 2 is related to the LTCD phase transition coming closer to the STCB phase transition. In terms

of ω, for the random case, the LTCD phase transition is almost over when the STCB phase transition begins, for the shuffled case there is more overlap between the high-ω part of the former and the low-ω part of the latter, while for the empirical case there is an almost perfect overlap between the two. The empirical behavior is illustrated by Fig. 3: within the $\omega \in [0.2, 0.4]$ interval, the decrease in LTCD is systematically accompanied by an increase in STCB. If one accepts that real-world systems are favorable for both LTCD and STCB and that the respective quantities used here are defined in a sensible way, this reasoning suggests that real-world systems function close to criticality, from the perspective of both measures: only at criticality or close to it are both quantities allowed to attain non-vanishing values in the empirical case. In order to stay away from criticality, the system would need
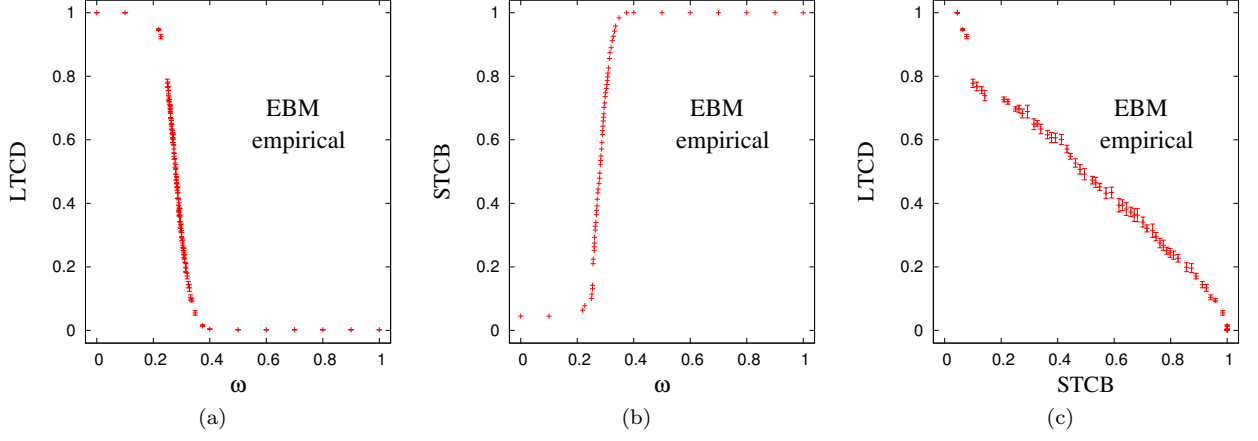
FIG. 3. The interplay between long-term cultural diversity and of short-term collective behavior for an empirical set of cultural vectors. Showing the LTDC(ω) dependence (a), the STCB(ω) dependence (b) and the ω-induced LTDC-STCB correspondence (c), for an empirical set of cultural vectors of $N = 500$, constructed from the Eurobarometer (EBM) data set.
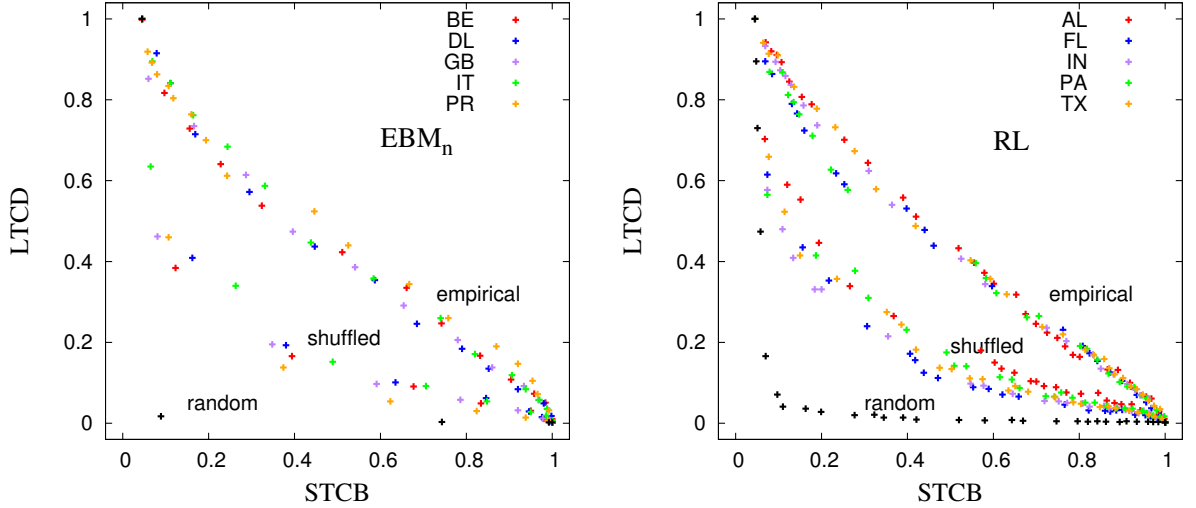

FIG. 4. The correspondence between long-term cultural diversity (LTCD) and short-term collective behavior (STCB) for empirical and shuffled sets of cultural vectors constructed from country-level and state-level samples of Eurobarometer-nominal (EBM$_n$) data (left) and Religious Landscape (RL) data (right) respectively. There are $N = 500$ elements in each set of cultural vectors. For visual clarity, error bars are omitted and the same colors are used for both the empirical and shuffled cases, while the LTCD-STCB curve is also shown for one random set of cultural vectors in each case.

to abandon either the propensity towards LTCD or the propensity to STCB. This suggests, as a speculation or conjecture, that the concept of self-organized criticality [36] might actually play an important role in describing trait formation and the interplay between exogenous and endogenous dynamics of traits. If this is correct, then a complete theory of cultural dynamics should have no need of fine-tuning the ω parameter.

Another important aspect is the robustness of the LTCD-STCB curves of Fig. 2 when switching from one geographical region to another, which is illustrated here by Fig. 4. This is done by focusing on the two data sets

which allow for division of the sample in terms of geographical regions, namely the Eurobarometer and the Religious Landscape. Moreover, only the nominal-variable information in the Eurobarometer is being used, for reducing the computational time required by the exogenous dynamics model, as well as for illustrating the robustness of the results with respect to the sample of cultural variables that are used. The empirical and shuffled LTCD-STCB curves are being shown for 5 EU countries (left) and for 5 US states (right) respectively. Only one random curve is shown, because, for a specific data set, the country/state-level SCVs are defined with respect to
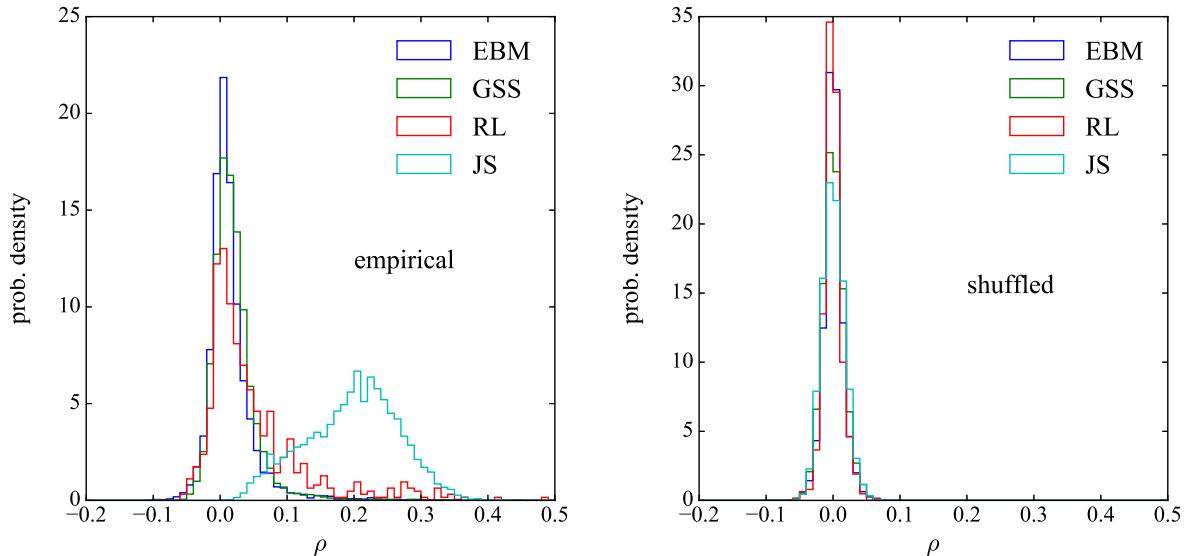
FIG. 5. Distribution of feature-feature correlation ρ for the empirical (left) and shuffled (right) versions of each of the four data sets (legend). Each histogram is normalized such that its integral is equal to 1, after being initially filled with $F(F-1)/2$ entries, where F is the number of features in the respective data set, each entry corresponding to one pair (k,l) of distinct features. For the normalization, the integral multiplies the bin content with the bin width $\delta\rho$ (the same for all histograms): the ordinate value of each bin is its relative frequency multiplied by a factor of $1/\delta\rho$.

the same cultural space, which is fully determined by the types and ranges of variables in the empirical data, which are the same regardless of the sample of people. Note that, for both data sets, the empirical and shuffled curves fall into clearly distinguishable bands. The empirical curves are systematically above the shuffled ones, while being again close to the second diagonal. This confirms the geographical universality of the structural properties inherent in empirical data.

When confronted with these results, one thinks of unavoidable similarities between questions in the survey, which induce correlations between cultural features. Since these correlations are destroyed by the shuffling procedure, it is tempting invoke them as an explanation for the discrepancy between an empirical LTCD-STCB curve and its shuffled counterpart. However, there is no reason to believe that such similarities are equally present in different empirical data sets, or that they are similarly distributed among the pairs of questions in the data set, since different data sets rely on completely different sets of variables. In fact, the measured feature-feature correlations $\rho^{k,l}$, defined via Eq. (3) are quite different across the four data sets used here. This is illustrated by Fig. 5, which shows how the values of these correlations are distributed for the different empirical SCVs (left), while also showing, for comparison, the distributions for their shuffled counterparts (right), which, as expected, are strongly peaked around 0 (the empirical and shuffled correlation matrices are shown in Figs. 6 and 7 of Appendix Sec. B). The departure of

the empirical distribution from its shuffled counterpart is clearly different across data sets, whereas the departure of the empirical LTCD-STCB curve from its shuffled counterpart is very similar across data sets, as shown in Fig. 2. Moreover, feature-feature correlations are typically small, given that any $\rho_{k,l}$ can take values within the $[-1,1]$ interval. These are indications that the properties captured by the LTCD-STCB plot are not (or not exclusively) due to feature-feature correlations, and that additional information destroyed by shuffling (including higher-order correlations) plays an important role. Such considerations enforce the belief that the observed properties are due to a more subtle, dynamical and universal mechanism.

V. DISCUSSION

This section discusses at length various points that have been open throughout the study, while mentioning possible directions for future research. First, Sec. V A further clarifies, at a rather philosophical level, the greater context and purpose of this research, while elaborating on concepts presented at the beginning of in Sec. I, mainly revolving around the contrast between endogenous and exogenous aspects of trait formation. Second, Sec. V B elaborates on certain assumptions, implications and possible weaknesses of the analysis conducted here.

A. Endogenous vs exogenous trait dynamics

In order to explain the problem of cultural trait formation and how this relates to the existing cultural dynamics paradigm, it was useful to distinguish, already at the beginning of Sec. I, between exogenous and endogenous mechanisms of trait dynamics. As explained, only the exogenous contribution is usually considered within the cultural dynamics paradigm, on which this work is largely based. In this sense, the focus is on social influence interactions taking place between agents, which in turn are only capable of changing their traits in response to the interactions. This type of description is attractive from a physics and a complex systems perspective, as it formulates macroscopic behavior in terms of interactions between a large number of microscopic constituents. Moreover, this approach also seems promising from a social science perspective, as it embraces both the dynamical nature and the multidimensional nature of culture, thus providing the ground on which a general theory of trait dynamics might one day be built, a theory which in principle would constitute a solution for the problem of preference formation. However, a complete theory should also account for endogenous processes, namely those that can take place in the minds of individuals even in the absence of social influence interactions. Formally, any kind of agent-level dynamical rules can be imagined (such as the cross-topic and reality-based consistency mechanisms described in Sec. I) implemented and studied (at least numerically), since agent-based modeling is highly versatile. The difficulty consists in empirically validating such rules and tuning the associated free parameters, since processes taking place in the human mind are very hard to probe experimentally.

Adding endogenous processes to exogenous dynamics models in a realistic way would definitely make good use of input from psychology and neuroscience, disciplines which are much more focused on the individual human mind and brain, respectively. Of course, by making use exclusively of notions from these two disciplines, one might fall in the other extreme, that of neglecting the exogenous aspect of trait dynamics. This is why an interdisciplinary approach is best suited for developing a general, complete theory of trait dynamics and for understanding how preferences are formed. This is something that was not attempted in this study, although it is a meaningful and arguably ambitious goal for future research. The purpose of this research was to indirectly study the combination between endogenous and exogenous processes taking place in the real world, by looking at isolated snapshots of this dynamics – the empirical cultural states.

One may wonder how important endogenous mechanisms actually are for trait dynamics, and whether they are worth mentioning at all. In a certain sense, this study suggests, together with Refs. [18] and [19], that the endogenous contribution is relatively high: had the structure present in empirical cultural states been dominated by the exogenous component, an exogenous dynamics model would have given similar results when acting on uniformly random initial conditions compared to empirical initial conditions. This expectation is natural if one believes that an empirical cultural state is somehow intermediate between an initial and a final state of exogenous dynamics, especially since uniform randomness, typically used for generating the initial state, can be interpreted as a complete lack of structure. But this scenario is obviously invalid, since exogenous dynamics clearly discriminates empirical cultural states from those generated in a uniformly random way. Thus, the results of this work confirm that empirical cultural states should be generally regarded as snapshots of a joint, endogenous-exogenous dynamical process. As discussed in Sec. V B, there is one argument against this reasoning, having to do with the fact that an empirical cultural state is obtained by sampling individuals from a geographically distributed population. However, based on previous results, one can argue that this in itself would not account for the discrepancy between empirical and random data.

It is worth noting that there are models of exogenous dynamics with so-called "cultural drift" [7], which involves arbitrary changes of traits irrespective of social influence. This can be seen as a first step in the direction of adding endogenous processes, if one interprets cultural drift as a primitive model of the endogenous component. However, cultural drift has been modeled as uniformly random, which, based on current results, apparently is not a good way of accounting for endogenous processes. In order to model these processes in a realistic way, more needs to be understood about these processes in the real world.

B. Empirical cultural states

Very important for this study is the possibility of converting empirical data from social surveys into sets of cultural vectors. This is a problem of data formatting that is challenging in several ways (see Sec. A). To a great extent, the task is to select a meaningful set of variables from those used in the survey that can be consistently interpreted as cultural features, in the sense of the Axelrod model. Specifically, one is interested in variables that record attributes of individuals that can be conceivably altered by social influence. In addition to demographic variables, which should be obviously disregarded, there are certain variables that need to be discarded for a variety of reasons (see Sec. A), even if many of them do encode culturally-relevant information. This amounts to an overall "loss of resolution" on cultural distances, which is not of much concern, as long as enough variables are left after this filtering: one can always argue that the survey could have contained more questions than it did, so there is always room for improving the values of the cultural distances. Indeed, the universal patterns shown in Sec. IV are visible despite the large variability in the

numbers of features constructed from the four data sets, and they are robust to reducing the number of features within one data set, which is explicitly shown in the case of the Eurobarometer: the LTCD-STCB analysis of Fig. 4 is carried out only using the nominal features, while that of Fig. 2 is carried out with both nominal and ordinal features, but leads to very similar results.

Once the formatting of the empirical data is complete, the resulting cultural state is assumed to be informative as a snapshot of the evolution of culture in the respective region of the world. However, in addition to exogenous and endogenous processes of trait dynamics, the empirical cultural states also reflects other effects, having to do with how the data is collected in the first place. The sampling bias is one such effect: certain people are less willing to participate in a social survey, or are less reachable by the survey because of living in a remote area. This bias somewhat alters the way individuals are distributed in cultural space, because their recorded preferences and opinions are likely correlated, to a certain extent, with their willingness to participate in a survey and to their place of residence. It is generally difficult to estimate how large this bias is and to compensate for it. Nonetheless, it is reasonable to expect that the strength and nature of this bias depends on the data set, because different surveys are technically conducted in different ways, for different purposes and in different geographical regions. The fact that the LTCD-STCB plot is independent of the data set and of the geographical regions (Fig. 4) suggests that the sampling bias is negligible for the purpose of this analysis, and that results discussed here cannot be and artifact of it.

Moreover, each empirical cultural state also reflects the fact that the set of selected individuals is a sparse sub-sampling from a geographically distributed population, of a much larger size. It is unlikely (and sometimes deliberately avoided through the survey design) that any of these individuals significantly interact directly with each other in the real world. However, when the exogenous dynamics model used here is applied to the empirical cultural state, one conceptually puts all these individuals in a "isolated container", where they can interact with each other for a long period of time, while constantly "mixing the content" of the container – since there is effectively no spatial or social structure constraining the way the agents in this model interact, while the interacting pair of individuals is selected at random at every moment in time.

This observation can be used as an argument for minimizing the importance of endogenous processes: one can argue that the structure inherent in an empirical cultural space is just the result of exogenous processes taking place in the real world, in combination with the sparse sub-sampling. Following this reasoning, one would expect that empirical-like structure can be obtained via the following procedure: run an exogenous dynamics model with a large population of agents that are geographically distributed (in physical space); define a set of cultural

vectors by selecting a small fraction of random agents in the final state; apply the LTCD-STCB analysis on this set of cultural vectors, for its shuffled and for its random counterparts, expecting that the universal behavior emphasized in Sec. IV is visible.

The authors' expectation is that this argument is not valid: would it be valid, one should also expect that the respective, more realistic model of exogenous dynamics would behave independently of the initial cultural state, unlike the minimalistic one used here. But this expectation already seems to be falsified, for small agent populations, by the results of Ref. [19], where such a realistic, Axelrod-type model is applied to empirical, shuffled and random initial cultural states, leading to different results for the three cases[2] – the respective version of the Axelrod model, which is an adaptation of that developed in Ref. [13], accounts for a physical, 2-dimensional space, for movement of agents in this space and for a dynamic social networks, with appropriate couplings to the dynamics in cultural space. Nonetheless, it is worth checking explicitly, in future research, the validity of such an argument, using the above procedure. If it does not work, the same procedure can be employed for a-posteriori validating models of trait dynamics with endogenous ingredients added to the exogenous component – see Sec. V A.

The fact that other, more realistic versions of the Axelrod model can be conceived apparently provides an argument against part of the claimed meaning and implications of present results: had the LTCD quantity been computed using a different model, one that would account for the physical space and for the social network, the dependence of the LTCD quantity on the bounded-confidence threshold ω would probably be different. This difference would likely manifest through a different position and width of the LTDC(ω) phase transition, which for empirical data would likely no longer overlap so well with the STCB(ω) phase transition. This reasoning suggests that the criticality claim made in Sec. IV is sensitive to details of the exogenous dynamics model used for computing the LTCD quantity.

However, this argumentation is misleading. In the present study, there is a certain parallelism between how the LTCD and the STCB quantities are computed as functions of ω – see Sec. III: they both condition social influence only on the distance in cultural space, via the ω threshold, while both disregarding the physical space and the social network. If the latter ingredients are to be included in the computation of LTCD, they should also be included in that of STCB, for the purpose of comparing the two phase transitions and for plotting the LTCD-STCB curves. For a certain value of ω, the presence of physical space and social network would increase the number of cultural domains in the final state of the Axelrod-type model, thus increasing the LTCD

—————

[2] although using an analysis tool somewhat different than the LTCD-STCB combination used here.

value, but at the same time would also increase the number of connected components in the initial state, thus decreasing the STCB value – the STCB quantity would essentially work with the connected component in the intersected graph, where a link between two agents is present when they are connected not only culturally, but also physically and socially; even if there is one connected component in each of the three graphs, one can still have multiple connected components in the intersected graph. Thus, it is reasonable to expect that the two phase transitions would shift in the same way upon adding model ingredients, if this is done consistently on both sides, while retaining the shape of the LTCD-STCB curve and justifying the criticality claim in a more generic sense. Nonetheless, this is another aspect worth checking explicitly in further research.

Yet another objection can be raised. For either data set, shuffling destroys correlations between cultural features. Thus, one can argue that the difference between the empirical and the shuffled regime of the LTCD-STCB analysis may simply be due to the presence of feature-feature correlations, which in turn are supposedly due to "design details" of the social survey, having to do with certain questions being similar to each other. Consequently, there would be no need to think about dynamical mechanisms responsible for the empirical structure, about integrating endogeneous and exogenous dynamical ingredients or about a complete theory of cultural dynamics and preference formation. However, the a-priori expectation is that design-induced correlations are relatively weak: collecting social survey data is expensive, so the survey should be designed such that it captures as much as possible of the relevant degrees of freedom, by minimizing the similarities among questions. Moreover, remaining similarities should be specific to each data set, whereas the LTCD-STCB analysis gives highly similar results for different data sets (the universality claim). To better illustrate this counterargument, feature-feature correlations were measured in Sec. IV and explicitly shown to be specific to each social survey, which is compatible with the idea that they largely depend on "design details". [3] In truth, feature-feature correlations can be seen as one of several manifestations of a non-uniform cultural space distribution, which is certainly also affected by (arguably unavoidable) a-priori, survey-dependent similarities between features, but not in an essential way: the universality of the results of the LTCD-STCB analysis should be due to something universal about the shape of the cultural space distribution,

which in turn is likely due to non-trivial, universal laws governing the dynamics of culture in the real world. It is also worth noting that one cannot say to what extent a correlation between two features is caused by an a-priori similarity between the two questions and to what extent it arises dynamically due to the combined effect of endogeneous and exogenous processes taking place in the real world. One can even argue that trying to disentangle the a-priori contribution is entirely meaningless, partly because the questions themselves are formulated by humans who interact with each other and with society.

VI. SUMMARY AND CONCLUSIONS

This manuscript started with emphasizing the important problem of trait formation. It argued that solving this problem would be equivalent to formulating and validating a complete theory of trait dynamics, which would embrace the dynamical and multidimensional aspects of culture. Both these aspects are accounted for in the existing cultural dynamics paradigm. However, this paradigm mostly focuses on social influence, thus on exogenous mechanisms, typically disregarding any endogenous ones. On the other hand, previous work had shown that the outcome of exogenous dynamics strongly depends on the cultural state used for the initial conditions. Specifically, there is more long-term cultural diversity (LTCD) for a given level of short-term collective behavior (STCB) for an empirical cultural state than for shuffled or random ones that are generated in the same cultural space. This provided a quantitative tool for highlighting the structure in empirical sets of cultural vectors, while raising the question of whether results obtained with it are independent of the empirical data set.

The study focused on addressing this question. It showed that the behavior of LTCD as a function of STCB is strikingly similar across data sets, implying that empirical cultural states have non-trivial, universal properties, which apparently cannot be recognized in the matrix of feature-feature correlations. Moreover, it pointed out that a presence of both LTCD and STCB in real world systems implies that these systems function close to criticality with respect to both aspects. One can object to this claim (Sec. V B) by pointing out an apparent sensitivity on the assumptions implicit in the formulation of LTCD. However, the assumptions behind LTCD are in this study paralleled by similar assumptions behind STCB, so modifying the exact formulation of LTCD should only be done while consistently modifying the exact formulation of STCB. One can thus reasonably expect that criticality would turn out to be robust if further investigations in this directions are made.

The LTCD-STCB analysis makes clear that convergence of cultural vectors under the action of exogenous dynamics takes place differently when starting with empirical data than when starting with random data. This suggests (Sec. V A) that an empirical cultural state is not

---

[3] Interestingly, the largest level of correlation is visible for the Jester (JS) data set, which is certainly the least expensive to collect, since respondents provide their answers online, via an automated platform. Moreover, the second-largest level of correlation is present in the Religious Landscape (RL) data set, which is arguably the second-least expensive to collect, since it relies on telephone interviews, while the other two data sets rely on face-to-face interviews.

an intermediate state between a random cultural state and an absorbing state of an Axelrod-type model. In other words, empirical structure is arguably not a consequence of exogenous mechanisms alone: endogenous mechanisms should have a significant contribution. An objection against this claim can also be formulated (Sec. V B): one can argue that exogenous dynamics appears to take place differently when starting with empirical data because this is sampled from a very large, geographically distributed population which had already been subject to exogenous dynamics for a long time. If this were true, there would be no need for endogenous processes when explaining the structure in the empirical data. Although this alternative explanation is worth testing in future research, there already exist indirect evidence against it.

This study demonstrated that one can learn about cultural dynamics from empirical data, even if this data consists of single, isolated, partial snapshots of this dynamics. Moreover, it emphasized the importance of endogenous processes and of their interplay with exogenous processes: including plausible endogenous processes, which was been very recently attempted for the first time [16], might be crucial for making existing dynamical models of culture realistic. The LTCD-STCB plot used here can conceivably be employed as a tool for validating such models.

Appendix A: Empirical data formatting

This section explains various details concerning the formatting of the empirical data. As previously mentioned, four data sets were employed, each of which was collected by different entities, for different purposes and in different formats. In order for the analysis and modeling conducted here to be carried out consistently, the important information had to be extracted from each data set and expressed in one, unified format. Essentially, this format dictates that each data set has to provide a certain number of ordinal features and a certain number of nominal features, where each feature has a certain number of possible traits (the range q of the feature), and that the traits of every individual in the data set are recorded with respect to all these features. This unified format can be effectively thought of as a table of traits, where the rows correspond to the features and the columns correspond to the individuals. There are various challenges involved when converting the data into this format. It is worth explaining first the challenges that are more generic, namely relevant for several data sets and then moving to the challenges specific to each data set.

One of the difficulties consists in deciding, for each variable, whether it should be used as cultural feature or not. The following is a (not entirely exhaustive) list of types of variables which are worth mentioning in this regard:

- demographic variables, such as those encoding "age", "place of residence" or "ethnicity" are discarded, as they do not record subjective human traits;

- certain variables, that were not seen as demographic variables by the survey authors, are also discarded if they recorded information about something that is too much in the respondent's past, or about something that cannot be easily related to subjective preferences, opinion, values, beliefs or behavioral tendencies that can be conceivably altered via social influence in a reasonably easy way; often, the boundary between what is subjective and what is objective not clear; nonetheless, one can strive to make these decision consistently at the level of every data set, which is what was done here;

- there are questions that ask opinions with respect to something that is differently defined for different people in the survey, such as: "how satisfied you are about how the the economy of this country is going recently?" – if there are people from different countries in the data set, or "how satisfied you are with your life?"; these questions are also discarded;

- questions asking the respondent to self-evaluate a certain, personal trait, such as "would you say about yourself that you are more conservatory or liberal on political affairs", are retained, assuming that the respondent mostly self-evaluates, in a reasonably objective way, a personal (subjective) trait, rather than expressing a subjective opinion about a the personal trait;

- certain variables containing relevant information are also discarded if, due to the survey format, they can only be answered when certain answers are given to other variables, or if the set of possible answers explicitly depends on answers given to other questions, regardless of whether these "other" variables themselves are selected or not; including such variables would introduce inconsistencies in the encoding of cultural vectors, the definition of cultural distance and the shuffling and randomization procedures.

The variables that are retained for further analysis need to be encoded either as nominal or ordinal cultural features. Deciding between the two encoding options was done here using the following criterion: if there are more

than two possible answers that are not "neutral" (see next paragraph) and they can all be conceivably ordered along the real axis, then the variable is encoded as ordinal; if, instead, there are only two answers (typically "Yes" and "No") in addition to the neutral ones, or if the non-neutral answers cannot be ordered along the real axis in a consistent way, then the variable is encoded as nominal.

Most variables retained from the data sets also allow for one or more "neutral" answers (often called "missing values" in social science research, although this term usually is somewhat more general). These are usually labeled as "Don't know", "Refused" or "Not Answered". For further analysis, these neutral answers are merged (if more than one are present). If the variable is to be encoded as nominal, neutral answers are mapped to one, additional cultural trait, side-by-side with traits originating from non-neutral answers. If the variable is to be encoded as ordinal, they are mapped to the middle of the ordinal scale – if there is an even number of possible answers, for each person, the choice is randomly made between the two answers closest to the middle of the scale.

Note that some data sets (GSS and EBM below) formally allow for another type of answer, labeled as "IAP" or "INAP" (inapplicable), which is here regarded as separate from neutral answers (although in social science research they are often all placed under the "missing values" umbrella term). IAP values are recorded, for certain respondents, when answers to a specific question are not expected from those respondents, for reasons having to do with the design of the survey. This happens for question that are only asked conditionally on answers given before. However, as mentioned above, these conditional variables are anyway discarded. Similarly, IAP values are also recorded for questions that are only asked to a certain sub-sample of the people, although not being conditional on some other question, in which case those questions are either removed or, if the sub-sample is large enough, the formatting is restricted to it. Finally, IAP values are also recorded for split-ballot or split-form variables (see GSS and EBM explanations below), in which case specific procedures are followed, which effectively discard all IAP answers before further analysis. Thus, regardless of how exactly they occur, one does not need to map IAP answers to any trait, as they are all filtered out as a consequence of other formatting rules. Note that for the RL data set, although IAP answers are not explicitly mentioned anywhere, this could have been the case, since there are questions that are conditionally asked on other questions – instead of IAP answers, system-missing values are present in the SPSS file, typically marked by the "." dot character.

First, this study made use of the **Jester 2 (JS)** data set [35], which consists of online ratings of jokes collected between November 2006 and May 2009. There are around 1.7 million continuous ratings (on a scale from -10.00 to +10.00) of 150 jokes from 59,132 users. For most users however, of the 150 jokes, only 128 are pro-

vided as items to be rated, as the other 22 were eliminated at a certain point in time. For this study, each of the 128 items is converted into an ordinal feature with 7 traits (by splitting the $[-10, 10]$ interval into 7 bins of equal size, while assuming that everything falling within one bin constitutes the same answer). Moreover, only the 2916 users that had rated all items were retained for further analysis – although this introduces some bias in the sample, one can argue that it is desirable to focus on individuals that have rated everything, as this is an indication of commitment on the respondent's side.

Second, the research used the **Religious Landscape (RL)** data set [34], which consists of opinions and attitudes on various religious topics, but also on various political an social issues. This data was collected in 2007 via telephone interviews from all states of USA – this study only used the data obtained from the continental part of the USA (without Hawaii and Alaska). There are multiple questions asking about the religious affiliation of respondents, which were all discarded. This is partly based on the assumption that religious affiliation is closer to a demographic variable than to a feature that can be easily altered via social influence, partly based on the very large number of answers and the nested, hierarchical nature of how they are organized. For this study, 36 cultural features were constructed (18 nominal and 18 are ordinal), for a number of 35558 respondents.

Third, the research used the **Eurobarometer 38.1 (EBM)** data set [32], which consists of opinions on science, technology, environment and various EU political issues (mainly related to the open market and the economy). The data was collected during November 1992, from 12 countries of the EU, via face-to-face interviews. In this survey, there are several blocks of "coupled" variables which are all discarded: within each block, there are explicit internal constraints on how answers can be given (such as answering "yes" to at most 3 questions out of 8 that are available), which do not allow for a consistent encoding as a set of nominal or ordinal features.

Another challenge when formatting the EBM data set is posed by the split-ballot procedure: the sample of people is split into 2 ballots, and certain questions are asked in slightly different versions (small differences in formulation, answers listed in different orders etc.) to the two ballots, while both versions are present in the SPSS file for all individuals – for every respondent, an IAP answer is recorded for the version that is not used for that respondent. The most meaningful approach is to merge the two versions and eliminate all IAP answers – if both versions are kept, strong structural artifacts arise in the matrix of cultural distances [19]. Most of the split ballot variables are encoded as ordinal and have the same range (same number of non-neutral answers) in both versions, such that a one-to-one correspondence can be made, similarly to Ref. [19]. Some of them are still ordinal but have different ranges in the two versions. In all these cases, there is a difference of only one trait among the two versions, such that one range is an even number while the

other is odd. In this case, the odd version is kept for the merging, which guarantees the existence of a middle trait to which all neutral answers can be directly assigned. The non-neutral answers from the even version are mapped to the closest answers in the odd version, in terms of the distance from the lowest-value answer, assuming that the distance between the lowest-value and highest-value answers is the same in the two versions (consistent with the definition of cultural distance in Eq. (1)). There is one split ballot variable which is encoded as nominal, in which case the difference consists in a second question being asked for one of the ballots, which is simply discarded. After all the formatting, 144 cultural features are constructed from this data set (54 nominal and 90 ordinal), for a number of 13026 respondents.

Fourth, the study used the **General Social Survey (GSS)** data [33], collected during 1993 in the USA via face-to-face interviews. The overall scheme of how questions are asked to respondents is arguably more complicated than for the EBM data set. First, there is a split-form procedure involved, which is equivalent to what is called "split-ballot" in the case of EBM: the respondents are split into two groups, with certain questions being asked in two, slightly different versions. All these questions are ordinal and have the same ranges in the two forms; they are handled like in the case of EBM. Independently of the split-form procedure, there is another procedure called "split-ballot", which is methodologically somewhat different: the sample of respondents is split in 3 ballots (A,B,C), while some questions are only asked to 2 of the 3 ballots (A and B, B and C or A and C). This is handled by discarding the questions asked to only 2 of

the 3 ballots. Independently of the split-ballot and split-form procedures, there is a set of questions, also used within the International Social Survey Program (ISSP), which are not asked to a small fraction of respondents (49 out of 1608 respondents). This is handled by discarding the 49 people not exposed to the ISSP questions. All in all, 133 cultural features are constructed from the GSS data (8 nominal and 125 ordinal), for a number of 1559 respondents.

Appendix B: Feature-feature correlations

This section illustrates in detail the correlations between cultural features, computed according to Eq. (3). The feature-feature correlation matrices of the four empirical SCVs are shown in Fig. 6, while those of the four shuffled counterparts are shown in Fig. 7. The ordering or rows and columns is consistent with the actual ordering of questions in the four data sets. This leads to a partial block-diagonal aspect of the matrices associated to the Eurobarometer and Religious Landscape data sets, for which questions that deal with similar topics tend to appear next to each other. Note that, empirical correlations rarely show strong deviations from their shuffled counterparts. There is a clear discrepancy between the Eurobarometer correlation matrix shown here and that shown in the Supplementary Information of Ref. [18]. However, the current study used a different, much more rigorous procedure of formatting the empirical data.

[1] M. Thompson, R.J. Ellis, and A. Wildavsky. *Cultural Theory*. Westview Press, 1990.
[2] E. Fehr and K. Hoff. Introduction: Tastes, castes and culture: the influence of society on preferences. *The Economic Journal*, 121(556):F396–F412, 2011.
[3] A. Cohn, E. Fehr, and M. A. Marechal. Business culture and dishonesty in the banking industry. *Nature*, 516(7529):86–89, 2014.
[4] A. Cohn, J Engelmann, E. Fehr, and M. A. Marechal. Evidence for countercyclical risk aversion: An experiment with financial professionals. *American Economic Journal*, 105(2):860–85, 2015.
[5] C. Castellano, S. Fortunato, and L. Vittorio. Statistical physics of social dynamics. *Rev Mod Phys*, 81(2):591–646, 2009.
[6] R. Axelrod. The dissemination of culture: A model with local convergence and global polarization. *J Conflict Resolut*, 41(2):203–226, 1997.
[7] K Klemm, V. M. Eguiluz, R. Toral, and M. S. Miguel. Global culture: A noise-induced transition in finite systems. *Phys. Rev. E*, 67:045101, 2003.
[8] K Klemm, V. M. Eguiluz, R. Toral, and M. S. Miguel. Nonequilibrium transitions in complex networks: A model of social interaction. *Phys. Rev. E*, 67:026120, 2003.
[9] M. N. Kuperman. Cultural propagation on social networks. *Phys. Rev. E*, 73:046139, 2006.
[10] A. Flache and M. Macy. Local convergence and global diversity: The robustness of cultural homophily. *arXiv:physics/0701333*, 2007.
[11] J. C. González-Avella, M. G. Cosenza, and K. Tucci. Nonequilibrium transition induced by mass media in a model for social influence. *Phys. Rev. E*, 72:065102, Dec 2005.
[12] Damon Centola, Juan Carlos Gonzalez-Avella, Victor M. Eguilez, and Maxi San Miguel. Homophily, cultural drift, and the co-evolution of cultural groups. *J. Conflict Resolut.*, 51:905, 2007.
[13] Jens Pfau, Michael Kirley, and Yoshihisa Kashima. The co-evolution of cultures, social network communities, and agent locations in an extension of Axelrod's model of cultural dissemination . *Physica A: Statistical Mechanics and its Applications*, 392(2):381–391, 2013.
[14] F. Battiston, V. Nicosia, V. Latora, and M. S. Miguel. Robust multiculturality emerges from layered social influence. *arXiv:1606.05641*, 2016.
[15] Alex Stivala, Yoshihisa Kashima, and Michael Kirley. Culture and cooperation in a spatial public goods game.
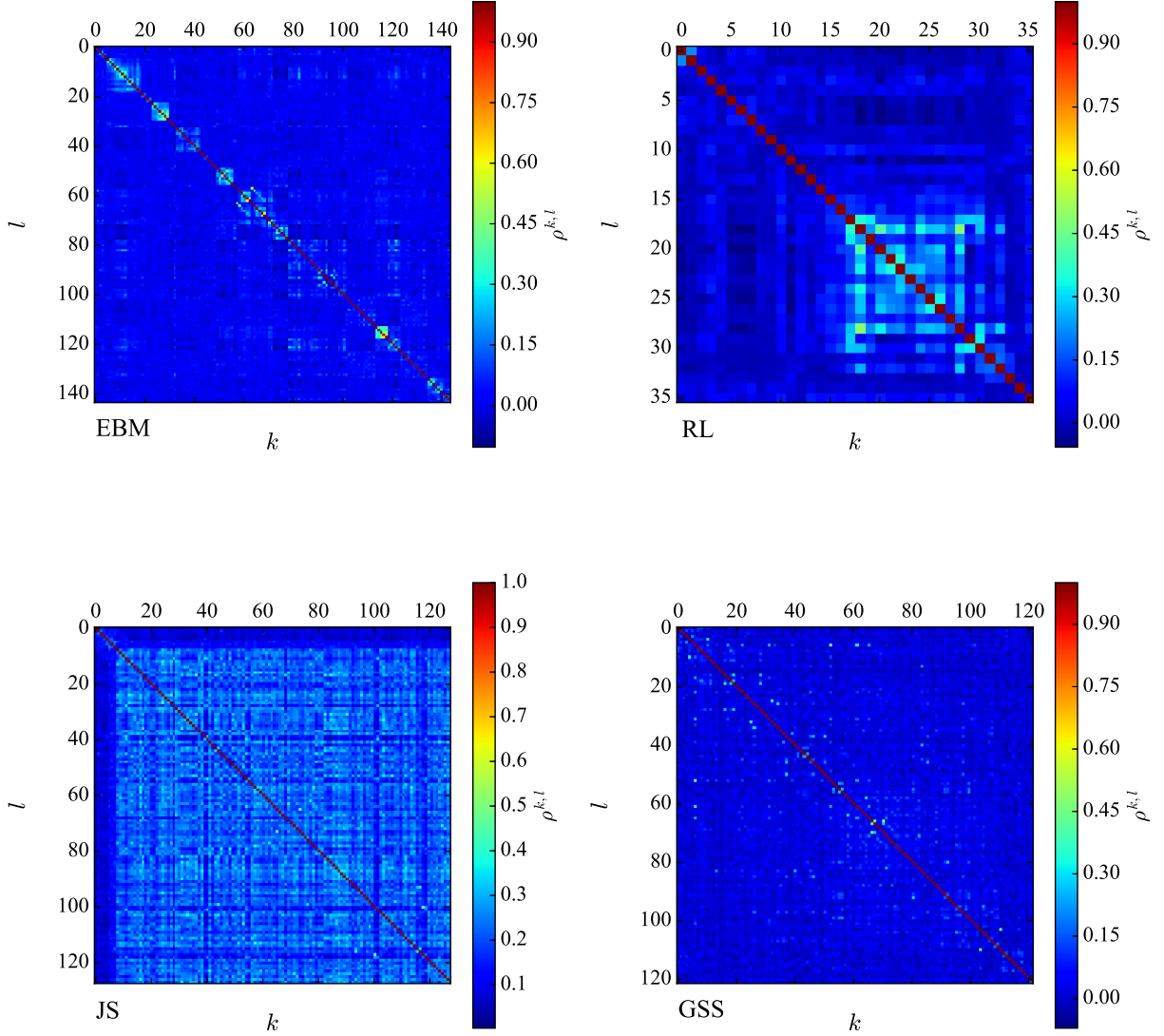
FIG. 6. Matrix of feature-feature correlations in empirical sets of $N = 500$ cultural vectors obtained from the four sources: Eurobarometer (EBM), Religious Landscape (RL), Jester (JS) and General Social Survey (GSS). Each grid point shows the correlation $\rho^{k,l}$ between cultural features k and l.

Phys. Rev. E, 94:032303, Sep 2016.

[16] Noah E. Friedkin, Anton V. Proskurnikov, Roberto Tempo, and Sergey E. Parsegov. Network science on belief system dynamics under logic constraints. *Science*, 354(6310):321–326, 2016.

[17] C. Castellano, M Marsili, and A. Vespignani. Nonequilibrium phase transition in a model for social influence. *Phys. Rev. Lett.*, 85:3536–3539, 2000.

[18] L. Valori, F. Picciolo, A. Allansdottir, and D. Garlaschelli. Reconciling long-term cultural diversity and short-term collective social behavior. *Proc Natl Acad Sci*, 109(4):1068–1073, 2012.

[19] A. Stivala, G. Robins, Y. Kashima, and M. Kirley. Ultrametric distribution of culture vectors in an extended Axelrod model of cultural dissemination. *Sci Rep*, 4(4870), 2014.

[20] R. Cont and J. P. Bouchaud. Herd behavior and aggregate fluctuatins in financial markets. *Macroecon Dyn*, 4(2000):170–196, 2000.

[21] J. A. Castner. Measures of cognitive distance and diversity. http://ssrn.com/abstract=2477484, 2014.

[22] M. Sherif and C.I. Hovland. *Social judgment: Assimilation and Contrast Effects in Communication and Attitude Change*. Yale University Press, New Haven, CT, 1961.
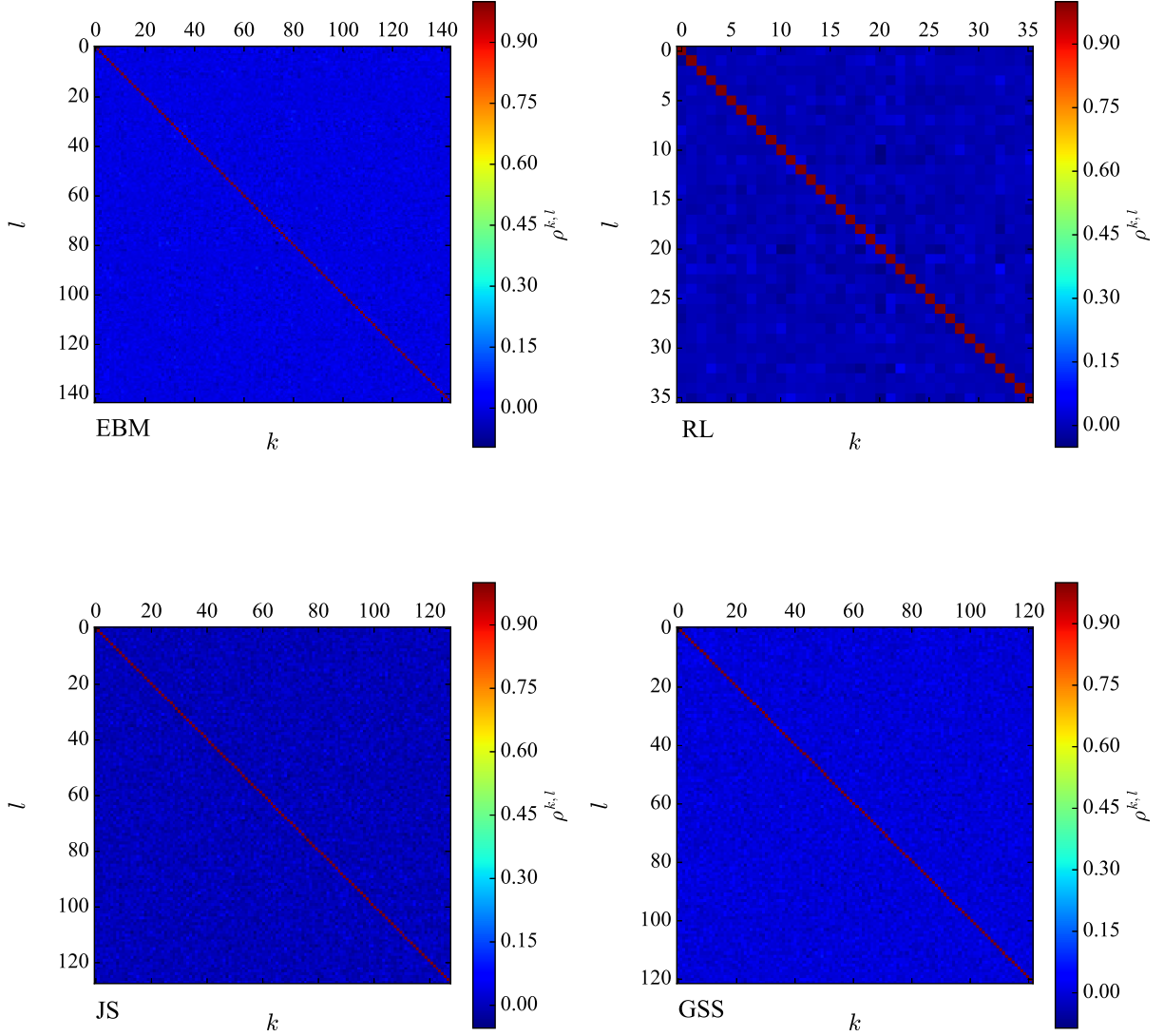
FIG. 7.   Matrix of feature-feature correlations in shuffled sets of $N = 500$ cultural vectors corresponding to the four empirical sources: Eurobarometer (EBM), Religious Landscape (RL), Jester (JS) and General Social Survey (GSS). Each grid point shows the correlation $\rho^{k,l}$ between cultural features k and l.

[23] J. Lorenz. Continuous opinion dynamics under bounded confidence: a survey. *International Journal of Modern Physics C*, 18(12):1819–1838, 2007.

[24] F. Barth. *Ethnic Groups and Boundaries*. Little, Brown and Company, Boston, 1969.

[25] R. Boyd and J. Richardson. *The Origin and Evolution of Cultures*. Oxford University Press, New York, 2005.

[26] E. Durkheim. *The Elementary Forms of Religious Life*. The Free Press, New York, 1912.

[27] J. P. Onnela and F. Reed-Tsochas. Spontaneous emergence of social influence in online systems. *Proc Natl Acad Sci*, 107(43):18375–18380, 2010.

[28] J. Ratkiewicz, S. Fortunato, A. Flammini, F. Menczer, and A. Vespignani. Characterizing and modeling the dynamics of online popularity. *Phys Rev Lett*, 105(158701), 2010.

[29] S. Fortunato and C. Castellano. Scaling and universality in proportional elections. *Phys Rev Lett*, 99(138701), 2007.

[30] B. Chakrabarti, A. Chakrabarti, and A. Chaterjee. *Econophysics and Sociophysics: Trends and Perspectives*. Wiley-VCH Verlag GmbH, 2006.

[31] S. Sinha, A. Chatterjee, A. Chakraborti, and B. Chakraborti. *Econophysics: An introduction*. Wiley-VCH Verlag GmbH, 2010.

[32] Karlheinz Reif and Anna Melich. Euro-barometer 38.1: Consumer protection and perceptions of science and technology, november 1992. `http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/06045`, 1995.

[33] Tom W. Smith, Peter Marsden, Michael Hout, and Jibum Kim. General social surveys, 1993 ed. `http://gss.norc.org/get-the-data/spss`, 1972-2012.

[34] Luis Lugo, Sandra Stencel, John Green, and Gregory et al. Smith. U.s. religious landscape survey. religious beliefs and practices: Diverse and politically relevant. `http://www.pewforum.org/2008/06/01/`, 2008.

[35] Ken Goldberg, Theresa Roeder, Dhruv Gupta, and Chris Perkins. Eigentaste: A constant time collaborative filtering algorithm. *Information Retrieval*, 4(2):133–151, 2001.

[36] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality: An explanation of the $1/f$ noise. *Phys. Rev. Lett.*, 59:381–384, Jul 1987.